

Hamline University DigitalCommons@Hamline

School of Education Student Capstone Theses and
Dissertations

School of Education

Summer 8-7-2015

Effective Teacher Performance Evaluation Systems and the Twin Cities Metro Area: Are local public schools using research-based valued practices?

Jeffrey Philip Escen Holtz
Hamline University, jholtz04@hamline.edu

Follow this and additional works at: https://digitalcommons.hamline.edu/hse_all



Part of the [Education Commons](#)

Recommended Citation

Holtz, Jeffrey Philip Escen, "Effective Teacher Performance Evaluation Systems and the Twin Cities Metro Area: Are local public schools using research-based valued practices?" (2015). *School of Education Student Capstone Theses and Dissertations*. 185.
https://digitalcommons.hamline.edu/hse_all/185

This Thesis is brought to you for free and open access by the School of Education at DigitalCommons@Hamline. It has been accepted for inclusion in School of Education Student Capstone Theses and Dissertations by an authorized administrator of DigitalCommons@Hamline. For more information, please contact digitalcommons@hamline.edu, lterveer01@hamline.edu.

Summer 8-7-2015

Effective Teacher Performance Evaluation Systems and the Twin Cities Metro Area: Are local public schools using research-based valued practices?

Jeffrey Philip Escen Holtz
Hamline University, jholtz04@hamline.edu

Follow this and additional works at: http://digitalcommons.hamline.edu/hse_all



Part of the [Education Commons](#)

Recommended Citation

Holtz, Jeffrey Philip Escen, "Effective Teacher Performance Evaluation Systems and the Twin Cities Metro Area: Are local public schools using research-based valued practices?" (2015). *School of Education Student Capstones and Dissertations*. Paper 185.

This Thesis is brought to you for free and open access by the School of Education at DigitalCommons@Hamline. It has been accepted for inclusion in School of Education Student Capstones and Dissertations by an authorized administrator of DigitalCommons@Hamline. For more information, please contact jneilson01@hamline.edu.

EFFECTIVE TEACHER PERFORMANCE EVALUATION SYSTEMS AND THE
TWIN CITIES METRO AREA: ARE LOCAL PUBLIC SCHOOLS USING
RESEARCH-BASED VALUED PRACTICES?

by

Jeffrey Philip Escen Holtz

A capstone submitted in partial fulfillments of the
requirements for the degree of Master of Arts in Teaching.

Hamline University

Saint Paul, Minnesota

August, 2015

Primary Advisor: Jennifer Carlson
Secondary Advisor: Jamie Holtz
Peer Reviewer: Colleen Holtz

ACKNOWLEDGEMENTS

Thank you to the hard work of the committee members who contributed greatly to development and completion of this capstone. I would also like to thank the school and district level administrators who worked with me and permitted the survey to be sent to their teachers.

TABLE OF CONTENTS

CHAPTER ONE INTRODUCTION.....	1
PURPOSE.....	3
PERSONAL PERSPECTIVE	5
BENEFITS OF CAPSTONE	6
CLOSING.....	6
CHAPTER TWO LITERATURE REVIEW	8
PURPOSE AND BENEFITS OF TPES.....	8
TPES & STATE LEGISLATURES.....	9
IMPROVEMENT OF TEACHING PRACTICE	10
SELF-REFLECTION.....	11
PROBLEMS AND RISKS ASSOCIATED WITH TPES	12
VAM USAGE.....	16
RESEARCH-BASED VALUED PRACTICES.....	19
PURPOSE OF IMPROVEMENT THROUGH FEEDBACK	20
THE EVALUATORS AND THE EVALUATIONS.....	22
VAM USAGE.....	23
KNOWN AND AGREED UPON EXPECTATIONS	25
VALUABLE FEEDBACK	26
CONCLUSION	28
CHAPTER THREE METHODS.....	30
RESEARCH SETTINGS AND SUBJECTS.....	30
THE PRAGMATIC WORLDVIEW	31
DESIGN AND RELEVANCE OF THE RESEARCH ACTION PLAN	32
THE SURVEY	33
<i>Table 1 – TPES Valued Practice Survey.....</i>	<i>34</i>
SUMMARY	39
CHAPTER FOUR RESEARCH RESULTS.....	40
SURVEY COLLECTION AND RESULTS.....	40
<i>Table 2 - Survey Results</i>	<i>41</i>
THE EVALUATORS	43
THE EVALUATION.....	44
VALUABLE FEEDBACK	45
RELATIONSHIPS WITHIN THE DATA	46
<i>Table 3 – Immediate Feedback & Experience.....</i>	<i>47</i>
<i>Table 4 – Immediate Feedback and Subject Matter</i>	<i>47</i>
<i>Table 5 – Multiple Feedback and Experience</i>	<i>48</i>
<i>Table 6 – Multiple Feedback and Subject Matter</i>	<i>49</i>
<i>Figure 1 - The evaluator has a deep understanding of effective classroom instruction.....</i>	<i>50</i>
<i>Figure 2 - The evaluator has a deep understanding of the curriculum I teach</i>	<i>50</i>

<i>Figure 3 - The evaluator has a deep understanding of effective classroom instruction.....</i>	<i>51</i>
<i>Figure 4 - The evaluator has a deep understanding of the curriculum I teach</i>	<i>51</i>
RESULTS CONCLUSION	52
CHAPTER FIVE CONCLUSION	54
LITERATURE REVIEW	54
IMPLICATIONS OF STUDY	56
LIMITATIONS OF STUDY	58
FUTURE RESEARCH	60
PERSONAL GROWTH & LEARNING.....	61
COMMUNICATION OF RESULTS.....	63
CONCLUSION	64
APPENDIX A - SURVEY REQUEST EMAIL (LETTER OF INFORMED CONSENT) ...	65
APPENDIX B - SURVEY COMPLETION CONFIRMATION STATEMENT	66
APPENDIX C – BREAKDOWN OF RESPONDENTS BY EXPERIENCE & SUBJECT MATTER.....	67
APPENDIX D – OPEN ENDED RESPONSES	68
REFERENCES.....	70

CHAPTER ONE

Introduction

The public education system in America has long been considered to be a key cornerstone of our society and its future success. One common viewpoint I often hear from stakeholders such as policymakers, parents, and voters is that schools are critical to long-term success because they prepare workers with the job skills for tomorrow's economy. I also hear the stance that public schools are a cornerstone of our society because they prepare our people to be educated voters who can participate in our democratic process. Another common viewpoint I hear about the importance of our public school system is that it is foundational to our future because it can function as a great equalizer for social mobility and economic status. I understand and appreciate all three of these perspectives due to my own personal experiences.

As a social studies teacher, one of the most important jobs I had was to emphasize civic responsibility and knowledge within my students. Having taught in a neighborhood where economic opportunities were scarce, I saw firsthand how our public schools empower citizens with the skills and knowledge that can allow them to move upward within American society. I also have worked outside of the public school system in areas where worker production and a company's success depends upon how well public schools have prepared employees for the workplace. Regardless of which perspective one finds to be more important, the fact remains

that many stakeholders view public education system as being of critical importance to America and its future.

Evidence of how important the public education system is for many American is abundant thanks to numerous public surveys on the topic. A 2012 Phi Delta Kappa/Gallup poll found that 89% of respondents thought it was very or somewhat important to try to close the achievement gap in America's public schools (Bushaw, 2012, p. 12). This same poll found that 97% of respondents felt it is very or fairly important to improve America's urban public schools. A CBS News Poll in 2010 found that 71% of the American public is in favor of using mandatory testing of students in public schools to determine how well the school is doing (CBS News Corporation, 2010). A quick search of the internet provides a random example of local level feelings towards America's public education system. A 2013 survey in Knox County, Tennessee asked respondents, on various factors, whether or not the factor was very important for improving the county's economy. An effective education was the number one factor with 90% of respondents defining it as "very important," (Bryant Research LLC, 2013). Additionally, respondents stated that teacher effectiveness was the most important factor, followed by parental involvement and poverty, which impacted the effectiveness of their schools. A national survey in 2013 of parents who had children in American public schools reaffirmed this when 96% of respondents stated that teacher quality was very or extremely important as a factor contributing to education quality (Tompson, Benz, & Agiesta, 2013, p. 4).

These snapshots of America show that stakeholders believe education is extremely important, they want to see improvement, and they understand that data regarding teacher quality and student performance are going to play major parts in the improvement process.

Purpose

As stakeholders demands that public schools do a great job of educating America's children and improve wherever possible, it could be argued that data needs to exist to show that improvement is occurring. Certain types of educational system data do exist already and can be accessed by interested stakeholders. As an example, in Minnesota one can go to the Minnesota Department of Education's website and pull up volumes of data involving the high school graduation rate of the state overall, of individual districts, and even of individual high schools (Minnesota Department of Education, 2014). Nationwide, one can review statistics that are collected through the National Assessment of Educational Progress (NAEP) program and see performance metrics of students in various subjects in certain grade levels (US Department of Education, 2014). These are just two methods in which an American taxpayer can look into the quality of schools by seeing how students at a macro level are performing.

However, while these and many other sources of data can provide information at the macro level about how students are doing on specific criteria, they may not provide information about how teachers are doing at the local level

and therefore many stakeholders are not getting the full picture of how schools are performing. As stated above, the many stakeholders feel very strongly that teacher quality correlates to the quality of education that is being provided within American public schools. We can gather and review data about graduation rates or standardized testing scores and have been doing this effectively for years. This data may allow a limited perspective to be made about how America's teachers are performing. What these data sets do not show, however, is how individual teachers are performing within their own classrooms. One possible way to show teacher performance is by evaluating teachers at the individual level.

Emphasizing the importance of quality teaching is not a new concept. Coleman, Campbell, Hobson, McPartland, Mood, and Weinfeld (1966) documented that characteristics of teachers explain more of the variance in student achievement and performance than any other measurable factor. Researchers in the 1980s found evidence that it not only is possible to start evaluating teacher quality but there can be verifiable benefits from doing so (Conley, 1987; Hiller, 1986; Zakariya, 1985). In the years since this research, programs created to evaluate teacher quality have started to migrate into America's public schools and progress is occurring in being able to report how teachers are performing. This project will not dive into the possible political and structural reasons as to why this migration and expansion is not complete or why there is opposition to it. However, the conclusions reached here hopefully can allow teacher evaluation programs to continue to expand by increasing our knowledge about how they work. I plan to do this by accomplishing

two core research goals. The first goal is to determine from the literature what the research-based valued practices are for effective teacher performance evaluation systems (TPES). The second goal is to determine if these practices are being used in evaluation systems that exist in Twin Cities metro area public schools.

Personal Perspective

I personally relate to the issue of evaluating teachers from the same perspective as some of the researchers noted above. My previous time spent teaching in the classroom was done at a school where an effective TPES existed. By seeing how the program could benefit staff, the administration, and the parents of students, I grew to truly appreciate the potential benefits that exist with TPES. After being in the classroom, I transitioned to a state agency where I currently am responsible for creating performance evaluation rubrics for over twenty different types of job positions. I have witnessed the difficulty with trying to evaluate work that many had previously said could not be measured or evaluated. I also have seen how an institution can quantify and report the performance of an institution's employees for growth and improvement purposes. My experience in this area, combined with having taught in a classroom that was using an effective TPES, is what leads and guides me in this capstone project. I do believe and understand that teaching is an extremely difficult occupation but that does not mean it cannot be evaluated. Various stakeholders want to know how students are performing but they also want to be assured that teachers are being evaluated and are being held to

the highest of standards. By establishing what the research-based valued practices are for creating, designing, and sustaining TPES programs at the local level, schools may be in a better position to meet the demand of knowing how teachers are performing.

Benefits of Capstone

Elected officials are responsible for enacting various public policies that impact our daily lives. Having lobbied at the Minnesota State Legislature and testified to multiple committees for specific policy issues, I have observed that policymakers prefer to have something to use as leverage or legitimacy to back up a reform effort. Policy reform efforts have the potential to impact various groups of people who, in their opposition to a reform, may use political capital to oppose the reform effort. By having an objective research document that supports a specific reform idea, policymakers may have a better chance of enacting a reform that may otherwise have been dropped in the face of political opposition. This project has the potential to add to the literature in a practical and real manner that could allow political leaders across the nation the additional leverage and legitimacy needed to continue implementation of effective teacher performance evaluation systems.

Closing

My goals in this capstone project are to 1) determine from the literature what the research-based valued practices are for effective TPES and 2) survey local Twin Cities teachers to determine the core components of the TPES they participate

in. This will allow me to answer the research question: *Are local Twin Cities metro area public schools using research-based valued practices in their teacher performance evaluation systems?* I plan to do this by analyzing a large volume of literature that has already been done on TPES and determine if researchers have found any common threads between programs that researchers have deemed effective. By conducting a review of prior research in this manner, I will be able to determine and document important valued practices that future administrators and policy makers can use when creating teacher performance evaluation systems for their own locales. Additionally, another aspect of this capstone project will be to survey teachers in locations where a TPES is in place. The intent of this is to gain local perspectives about what the structure currently is of TPES's locally, how teachers feel about their TPES, and if their TPES is following known research-based valued practices.

CHAPTER TWO

Literature Review

This literature review will lead the capstone project towards the goal of being able to answer the following question: *Are local Twin Cities metro area public schools using research-based valued practices in their teacher performance evaluation systems?* There are two steps that this capstone will take in order to answer this question. First, the literature review of this chapter will look into what the research-based valued practices are for effective teacher performance evaluation systems (TPES). Second, based upon these findings, this capstone will determine if these best practices are being followed and carried out locally in Twin Cities metro area public schools. This chapter describes several main themes that are apparent within the literature that are pertinent to this discussion. The first theme within the literature is what researchers advocate as being the purpose of TPES. The second major theme that exists in the literature is that problems and risks exist with being able to carry out TPES effectively. The final theme this chapter discusses is the compilation of valued practices that researchers have identified for developing and implementing effective TPES. These three themes combined represent the main findings of this literature review regarding teacher performance evaluation systems.

Purpose and Benefits of TPES

The first chapter of this capstone presented several reasons why teacher performance evaluation systems should exist and how they would be beneficial. The reasons listed varied from the idea that evaluations allow taxpayers the

assurance that taxes are well spent to seeing how evaluations allow teachers and administrators to improve teaching practices. When it comes to the literature on the TPES, the first main theme that is apparent involves the defined purpose of TPES.

TPES & State Legislatures

One of facets from this review was discovering that a defined purpose and benefit of TPES could be found outside of scholarly research in legislative state statutes. In a research paper that discusses various states in the union and how each have differently approached the policy area of collective bargaining and the evaluation of teachers, Paige (2013) discusses that Florida codified the purpose and benefit of TPES for its citizenry. Paige also mentions that, as per Florida Statute § 1012.34(1)(a), evaluating the performance of public school teachers in the state of Florida is done for the “purpose of increasing student learning growth by improving the quality of instruction” (p.12) in its schools. The statute itself is a recent addition and at the time of its passage had called for the development of procedures for instituting system wide teacher evaluation program. The statute later states that an additional purpose of their statewide TPES is to allow for data to be used when districts and schools develop internal improvement plans. While Paige gives the example of Florida explaining specifically in its own state statutes the purpose of its TPES, another codified example exists much closer to home in the state of Minnesota.

The Minnesota State Legislature has a law requiring that, by the fall of 2014, all school districts in the state will have to have started to implement a teacher performance evaluation system (Teachers and Other Educators, 2014). The statute required a newly created task force develop a model TPES that local districts could choose to use on their own. If a local district made the choice of not using the model developed by the task force, then the district would be required to develop a TPES on their own that conforms to several key requirements. Minnesota is an example of a state that, just like Florida, legally codified the purpose for mandating a TPES across the state: to “improve student learning and success,” (Teachers and Other Educators, 2014). It later states that a benefit and purpose of evaluating teachers is to support teachers and allow them to improve their practice. Both of these states provide examples of purposes and benefits of TPES. In addition to these states providing codified examples of the purpose and/or benefit of TPES, the scholarly research also provides several examples as well.

Improvement of Teaching Practice

One of the core purposes of TPES from the literature is to allow for the improvement of teaching practice (Chukwubikun, 2012; Conley, 1987). The basic premise is that TPES allows for a feedback that teachers can act upon. Feedback, whether it is based upon observation, test scores, another factor, or a combination of all of these, provides teachers information about their performance. If that feedback is consistent and regular, teachers are given the chance to know where

they are doing well and where they may need to try something new (Olivia, Mathers, & Laine, 2009).

One of the keys in the improvement process itself is that an observable problem is identified and communicated to the classroom teacher. The identification of a problem is thus the next purpose and benefit of TPES. Olivia (2009) recognizes that TPES should be used for both recognizing areas of excellence but also for identifying specific problems that can be and should be fixed (Olivia, Mathers, & Laine, 2009). It logically flows that if a core purpose and benefit of TPES is to allow for improvement of teaching and student performance, you first have to identify problematic areas.

Self-Reflection

Another benefit that the literature discusses is that TPES allows for teachers to be self-reflective (Taylor & Tyler, 2012). An example of this exists with the Teacher Evaluation System (TES) in the Cincinnati Public School System. Taylor and Tyler (2012) conducted an in-depth review of TES and, amongst their findings, they found evidence that no matter the criteria, providing feedback through the evaluation process caused an increase in self-reflection opportunities for teachers. In concluding that there were larger student gains that occurred within classrooms where TPES was used versus classrooms where it was not being used, Taylor and Tyler mention that providing feedback to teachers was of critical importance because it allowed teachers to either develop or change their teaching methods in a

manner that caused an improvement in student performance (Taylor & Tyler, 2012).

Overall, three core purposes/benefits arose from a thorough review of the pertinent literature involving teacher performance evaluation systems (TPES). The first such theme is that TPES is meant to improve the performance of student learning and growth. The second theme is that TPES is designed to help teachers improve upon their own teaching methods. The last and related theme is that TPES can improve teacher quality and thus student achievement by providing feedback to teachers about their performance in a manner that provides an opportunity for self-reflection. A central tenet that runs through these interrelated themes is that TPES is meant to create a flow of information about teacher and student performance that can then be acted upon so improvements can occur. While this would appear to be a worthwhile and noble cause, the literature also discusses that TPES is not without controversy or risk.

Problems and Risks Associated With TPES

The relevant research presents a picture of why problems have surrounded TPES and what the possible risks are that associate with its implementation. These possible risks include not using a qualified or trained classroom observer, having evaluations that lack feedback for improvement purposes, failing to connect evaluations to professional development opportunities, and the reliability and validity concerns of specific student testing data. One area the literature does

consider is how teachers who are being professionally evaluated feel about the evaluation process. Specifically, multiple researchers have surveyed teachers to determine how they feel about the people who are completing classroom observations as part of their overall evaluations. Stark and Lowther (1984) found in their large scale survey of teachers in Michigan that 85% were accepting of having an administrator perform the classroom observation and 73% were accepting of having a peer teacher perform the assessment. Both of these statistics show that the surveyed teachers were receptive towards having the observations be done by someone from within the school and the education profession. Olivia et al. (2009) found that teachers want to have an observer that has a deep understanding and knowledge of the curriculum. This is because of the feeling that teaching methods may vary and correlate to the type of curriculum one is teaching and therefore it would be best to have an observer who understands the curriculum that is being taught in each classroom. In their review of the Cincinnati teacher evaluation system (TES), Taylor and Tyler (2012) found the use of peer teachers for observations was advantageous because the observed teacher would be more receptive toward the feedback than had it been from an administrator. Further, a 2007 analysis of multiple teacher evaluation programs in the Midwest region found that, even though it is extremely important, less than 10% of the programs required training for staff who conducted classroom observations (Brandt, Mathers, Olivia, Brown-Sims, & Hess, 2007). This was found by the researchers to be problematic as

observers need to be trained so they are properly prepared to conduct a valid and non-biased observation of a teacher's performance.

All of these studies combine to create a controversial question: Who is qualified to be the person conducting classroom observations of teachers as part of a teacher performance evaluation? Should it be someone who is properly trained to be a qualified observer and evaluator, should it be someone who is highly knowledgeable about the curriculum being taught, and what preference should exist between using a teacher versus an administrator? There is a difference in preference amongst teachers as to whether they would prefer to have a peer teacher or an administrator be the observer and evaluator. While Taylor and Tyler describe a preference of using peer teachers instead of administrators as the observer, there also is a risk and cost that can arise from that. By using a peer teacher to be an assessor, one who is a high performing teacher in their own right, that teacher is now spending less time in the classroom with students and is likely being replaced by someone who is less experienced and less effective (Taylor & Tyler, 2012). This may place administrators in a very tough spot. While teachers may have a preference for a peer teacher as the classroom observer, the potential gains from helping other teachers improve may be negated by taking a high performing teacher out of the classroom. The alternative of using an administrator as the classroom observer for teacher evaluations may create controversy as well as concern that not all administrators may have a deep understanding of all curricula being taught within the school

As discussed earlier in this chapter that one of the purposes of teacher performance evaluation systems (TPES) is to provide feedback to teachers that they can reflect upon and then take the necessary actions to improve their own performance. A substantial problem that the literature discusses is that often the teacher being evaluated does not receive feedback on how he or she can improve. The danger of this is that “if teachers are not provided with clear signals about legitimate ways in which to improve their practice, there is the danger that teachers will focus instead on teaching test-taking skills at the cost of teaching other, more difficult to measure (but valuable) skills” (Kane, Taylor, Tyler, & Wooten, 2001, p.3). Weisberg (2009) takes a deeper look into this issue by analyzing the TPES used by 12 different school districts in four different states and by surveying approximately 15,000 teachers and 1,300 administrators in those districts (p.32). Over 73% of teachers stated that their evaluations made no commentary on how or where to improve their performance (Weisberg, Sexton, Mulhern, & Keeling, 2009, p.33). This may be because the evaluators felt that 73% of teachers did not need to improve. This also may be because evaluators not understanding the importance of providing critical feedback to teachers. Weisberg and his fellow researchers also mention that instead of evaluations being used to improve student and/or teacher performance, they found that teacher evaluations are nearly exclusively used for the purpose of making employment decisions such dismissal and remediation (Weisberg et al., 2009). This poses a significant problem for teachers in that the added benefit of feedback for improvement is not occurring and, instead, employment decisions such

as dismissal are occurring without an opportunity for teachers to improve. This also is a substantial risk for schools as TPES was not designed solely to be an information source for making employment decisions. If teachers are not given an opportunity to know where and how to improve based upon the evaluations themselves, then the schools are missing a substantial opportunity by using them only for making employment decisions.

VAM Usage

The other problem with TPES that the literature discusses involves growth data that is used to measure the effectiveness of a teacher. The common term for this type of data is “Value Added Measure,” or VAM. Corcoran’s plain language explanation for how to conceptualize VAM is to “think of a teachers’ value added as her students’ average test gain, properly adjusted for other influences on achievement” (Corcoran, 2010, p. 5). It essentially is an attempt to measure the impact that an individual teacher had on the growth of a student or a group of students. A plain language example of how VAM data can apply to a TPES program can involve a school that uses the Measures of Academic Progress (MAP) test for students three times per year. An example of this may be that during the first week of school, Mr. Johnson’s class of 25 third graders read with an average reading comprehension level of 2.8 but by the end of the year they were at 3.7. The measurement on Mr. Johnson’s evaluation would state that his students grew by 0.9

grade levels in reading comprehension during the school year and this growth would be entirely contributed to Mr. Johnson.

One of the problems that Corcoran identifies with the use of VAM in this manner is that it entirely depends upon a skill being able to be assessed in a standardized test. Because of this, not all subjects and therefore not all teachers can have a valid VAM that is attributed to them (Corcoran, 2010). In the case of using the MAP test to create a VAM for each teacher, this is indeed a limitation as not all subjects have a MAP test that could be used. There are for example, no MAP tests for any social studies subjects. Another problem that Corcoran describes is that, per his analysis of how New York City and Houston public schools used VAM as part of their TPES, there is a high level of variability from one year to the next for individual teachers. Corcoran describes this as a problem because a valid and reliable VAM should not have a large variance such as this as teachers do not vary in the quality of their performance significantly from one year to the next (Corcoran, 2010). Corcoran is not alone in voicing concerns over the potential problems associated with the use of VAM in TPES.

The premise of VAM rests upon the ability to statistically control all variables that impact the performance of a student besides that of his or her teacher. The reason for this is then a teacher can have a true measure of how much he or she contributed towards the growth of their students. Koedel (2009) found that one variable that VAM's are currently not controlling for is impact of multiple teachers on a single measured skill. Specifically, when it came to a subject such as reading

comprehension, multiple teachers in varied subjects have an impact on the growth of reading comprehension within a given sample of students but the VAM that measures that growth is only attributed to the reading teacher (Koedel, 2009). This first means that the reading teacher has a VAM that is not due solely to his or her own contributions. It also means that the contributions of other teachers are not being measured for feedback purposes either.

Another internal school variable that can impact the validity of VAM involves the quality of teaching within a school overall. Jackson and Bruegman (2009) mention that students in a specific class can perform better, and therefore impact a VAM in a positive manner for that classroom teacher, when that classroom teacher works with other highly effective teachers. Essentially this means that, just as Koedel found, multiple teachers can have an impact on the VAM for a specific classroom teacher. If a teacher of low quality teaches in a school with an abundance of high quality teachers, the VAM for the low quality teacher would potentially be artificially higher than what his or her actual VAM contribution really is.

Overall, the literature does present multiple problems and risks associated with TPES. Even though one of the goals of TPES is to provide feedback for teachers so they know where they can improve, researchers have found that often the evaluations themselves are void of any critical feedback that can be acted upon. Additionally, the literature describes that instead of evaluations being used for multiple purposes, they are often being used solely for making employment decisions. It also discusses specific limitations and problems associated with the

usage of VAM for teacher evaluations. While the literature does identify specific problems associated with certain parts of TPES, the literature also clearly describes what several research-based best practices are within effective teacher performance evaluation systems.

Research-Based Valued Practices

The literature presents a strong picture for what the defining components are of an effective and well-constructed teacher evaluation performance system (TPES). The six components that will be discussed in this section are not the only components that the literature discusses but they are the most common components that the literature review found as relating to effective TPES. The six components of an effective TPES program that this capstone will discuss are based upon a synthesis of the subsequent research:

1. The TPES has a purpose of increasing teacher and student performance through evaluation feedback
2. The TPES uses targeted classroom observations by peer and administrative evaluators
3. Classroom observations are done by trained and qualified evaluators
4. Proper and statistically controlled VAM data is used for limited purposes
5. Clear expectations exist for teachers within the evaluation that are based upon agreed criteria and goals

6. The TPES provides timely and frequent evaluation feedback that is linked directly to professional development to allow for targeted teacher growth

Purpose of Improvement through Feedback

The main purpose of TPES is to provide critical feedback to teachers so they and their students can improve. The literature describes that effective TPES needs to have this purpose be something that is engrained into its design and implementation. Weisberg et al. (2009) describes this purpose as an absolute and that the core purpose of TPES has to be to improve teacher growth and effectiveness. Olivia et al. (2009) discusses that effective TPES is designed for the purpose of identifying excellent teachers, locating areas where problems exist, and creating feedback that allows for targeted professional development and growth.

The most effective type of TPES is a system in which multiple types of measurements exist to evaluate teachers (Looney, 2011; Kane, Taylor, Tyler, & Wooten, 2011). A common type of measurement and the next valued practice of this capstone is the classroom observation. This involves an observer coming into a specific classroom and evaluating what he or she sees from the classroom teacher with a pre-determined rubric of expectations. How a classroom observation is designed and what that rubric looks like may vary from district to district (Hiller, 1986). According to Kane and his fellow researchers, one key to having an effective classroom observation is to have the rubric measure only skills and teaching practices that can acted upon and/or improved (Kane et al., 2011). The reason for

this is observing and reporting upon a skill that a teacher cannot change or improve is not an efficient use of resources and it negatively impacts the legitimacy of the observation. Hiller notes that the skills or teaching practices that a classroom observation rubric measures may be dependent upon whether or not there is a preferred teaching style within a school and if there are specific method goals that a school administration has established for teachers (Hiller, 1986). An example of this variance is how an observer evaluates a teacher who is dealing with a discipline issue. A school that has an established method and required action steps for resolving discipline issues would have the observer evaluate a teacher based upon those standards. Alternatively, a school that does not have a set standard for resolving discipline issues may not have that standard in the observation rubric and the teacher would not be evaluated in that area.

Additionally, multiple researchers note that effective classroom observations cannot be infrequent or for brief amounts of time (Chukwubikem, 2012; Conley, 1987; Mathers & Olivia, 2008). Plainly stated, “evaluators cannot accomplish this goal with a sample of only a few hours or observation or with an observation of only one class” (Chukwubikem, p.23, 2012). Thus the lower the frequency and the lower the amount of observed time, the greater the chance that the observation results are not reliable. They note that one evaluation of a teacher should be based upon at least four or five separate classroom observations. The issue of resources may determine how often a district can place an observer into a classroom but Mathers and Olivia (2008) and Chukwubikem (2012) strongly argue that infrequent or brief

observations may not even be worth it because of the potential harm to its reliability.

The Evaluators and the Evaluations

While the overall goal of providing feedback is for teacher and student improvement, the next components of effective TPES involve how the feedback comes to be. First, the individual evaluators who conduct classroom evaluations should be qualified, trained, and non-biased. Stark and Lowther (1984) note that historically, a building administrator such as a principal has conducted classroom evaluations. They also note that teachers have a slightly higher preference of having the evaluator be a peer teacher instead of an administrator, although overall both are accepted styles per the teachers that they surveyed. Regardless of whether the evaluator is a peer teacher or an administrator, a key requirement is that the evaluator has a deep knowledge of the curriculum, content, and instruction (Mathers & Olivia, 2008; Weisberg et al., 2009; Zakariya, 1985). A noted practice is to give the evaluator access to the lesson plans prior to the observation (Mathers & Olivia, 2008). As they explain, if evaluators do not have the lesson plan and student accommodations are needed in the lesson, “it would be difficult for the evaluator to know if these accommodations are implemented appropriately” (Mathers, p.5, 2008).

Another important research-based valued practice for the classroom evaluation is the evaluator needs to be trained in how to conduct the evaluation.

Chukwubikem (2012), Olivia, Mathers, and Laine (2009) and Conley (1987) note that an effective classroom evaluation requires having a trained evaluator. Training can ensure that each evaluator knows what the pre-designed rubric is designed to evaluate which can then reduce the amount of bias that can occur in the process. An effective classroom evaluation requires that the evaluator be prepared to conduct a valid and reliable evaluation. Stark and Lowther (1984) state that peers and school administrators can both be used and can perhaps even provide different types of feedback based upon their own skills sets as evaluators. Having both peer and administrator evaluations also aligns with the need to have multiple sources of measurements. Chukwubikem (2012) argues specifically that schools should “expand the number of people involved” (p. 560) with observations because expanding the number of participants expands the perspectives that the evaluation feedback offers to teachers. Besides classroom evaluation data and measurements, there also is another major type of measurement the literature discussed called value added measurement.

VAM Usage

As discussed previously, a value added measurement (VAM) is a statistical measurement that attempts to rate the level of contribution that a specific teacher had on his or her students. While VAM was discussed previously in this capstone as a potential problem area due to its limitations, it is still a measurement that is extremely valuable when used wisely for TPES. VAM provides another source of

data for evaluating teacher performance by attempting to quantify how a teacher impacts student growth. Maslow and Kelley (2012) point out that “linking evaluation to student test scores, as recommended by Race to the Top, could broaden the scope of information obtained through the evaluation system and provide a richer source of data for formative and systemic evaluation (p.629). They do acknowledge that because VAM data is not perfect, it should not be the only source of data that is used in TPES and instead should be combined with other sources of data. Kersting, Chen, and Stigler (2013) argue that VAM can be used as a valid and effective measurement tool for TPES. Regarding concerns over validity with VAM data, they describe that an individual teacher’s VAM should be based upon a sample size that is statistically large enough to reduce the standard error of the measurement. Their recommendation was that a VAM based upon at least 15 students would suffice but 20 would be preferred (Kersting et al., p.28, 2013). In countering claims that VAM’s are not an effective source of data because prior studies have shown large year-to-year variance in the VAM measurement of specific teachers, Kersting, Chen, and Stigler point out that the year-to-year variance they observed is not significantly lower than the level that researchers traditionally consider to be an acceptable year-to-year variance. Additionally, they state that the level of variance that researchers consider to be acceptable for research projects is artificially higher than what would be reasonable specifically for the VAM of a teacher. This is because change occurs from one year to the next for a teacher

whereas the acceptable level for researchers overall is based upon controlled experiments where all but one variable do not change.

Mathers and Olivia (p.9, 2008) describe that the main limitation of VAM is not that of validity or reliability but is that of acknowledging that not every aspect of teaching can be measured by a VAM:

“Those who teach social studies, physical education, music, art, special education, as well as K-2 teachers and many middle and high school teachers, cannot be assessed using student test scores because not all are assigned a defined set of students in a classroom and not all subjects are tested every year or in every subject.”

While a VAM may not be able to measure everything that a principal, a school board, or others may want to know about a specific teacher, Mathers and Olivia argue that a VAM can still be used and should be included as another measurement tool for TPES in areas where it is applicable. Having discussed measurements such as VAM and classroom observations as well as the importance of having effective observers, the next research-based valued practice does not involve measurements at all.

Known and Agreed Upon Expectations

In order to have an effective TPES, teachers must be aware of what is expected of them and they need to be involved in the evaluation process. The literature states that effective TPES requires that there is clear communication to teachers about the goals they will be evaluated on based upon the observation

rubric (Chukwubikem, 2012; Looney, 2011; Weisberg et al., 2009). The evaluation criteria and rubric need to be known to teachers prior to their evaluation.

Additionally, it should be noted that teachers are favorable towards being observed and assessed based upon set criteria and goals if they already had agreed to their use. “Teachers were quite favorable toward the idea of being assessed on their success in achieving objectives negotiated in advance” (Stark, p.102, 1984). By being aware of the observation criteria and by having agreed to their use, teachers then know what to expect and have an opportunity to ensure that they are carrying out their lesson plans in a manner that is consistent with the evaluation’s rubric. By having a TPES possess established and agreed upon observation criteria and goals through its rubric, the TPES may be in a position to effectively implement the next research-based valued practice that the literature discusses and the final one of this chapter: valuable feedback.

Valuable Feedback

Of the many components that effective teacher performance evaluation systems possess, an extremely important one is making sure that feedback from the evaluation is valuable to teachers (Delvaux, Vanhoof, Tuytens, Vekeman, Devos, & Petegem, 2013). There are several ways in which evaluation feedback to teachers can become valuable. First, feedback is valuable to a teacher if it is given in a timely manner (Looney, 2011). One suggestion for how this could occur is by setting up a conference meeting in which feedback is given to the teacher in a constructive

manner (Chukwubikem, 2012). This conference may involve the teacher and the administrator who compiled all of the pertinent data and observation feedback or it may also include peer teachers who may also been involved with conducting classroom observations.

Another way in which evaluation feedback becomes valuable for teachers is when it is connected to professional development opportunities. Because a core purpose of TPES is to allow for teacher improvement and thus potential improvement in student achievement, an effective TPES must tailor evaluation feedback to teachers in a manner that targeted professional development and growth opportunities can follow (Benedict, Thomas, Kimerling, & Leko, 2013; Chukwubikem, 2012; Looney, 2011; Mathers & Olivia, 2008; Weisberg et al, 2009). The literature also notes that an effective TPES connects evaluation feedback to other policies such as tenure approval, dismissal, salary or bonuses, and other intrinsic or extrinsic rewards that teachers would value (Chukwubikem, 2012; Stark & Lowther, 1984; Weisberg et al, 2009). By linking feedback and the results of an evaluation to things that matter to teachers such as money or employment status, the research shows that teachers then are more invested in the evaluation system and in achieving growth for themselves.

Because an effective TPES connects evaluation feedback and results to things that directly can positively or negatively impact a teacher, the feedback should be common enough in frequency so no teacher is ever caught off guard from a possible negative consequence (Weisberg et al, 2009). In a system with frequent and valid

feedback, “unsatisfactory ratings will not be anomalous, surprising, or without clear justification” (Weisberg, p.35, 2009). By providing multiple opportunities for useful evaluation feedback to teachers who may have room for improvement, they have a chance to prevent a negative consequence from occurring and instead can increase the chance of a positive consequence occurring. By providing valuable and effective feedback, an opportunity is created for instructional practices to be improved so schools can better meet the needs and demands of their students (Benedict et al, 2013). Feedback that is frequent, timely, and aligned with agreed upon goals provides teachers the chance to connect with professional development opportunities and improve their practice.

Conclusion

The goal of this capstone is to answer the following research question: *Are local Twin Cities metro area public schools using research-based valued practices in their teacher performance evaluation systems?* In order to accomplish that goal, a thorough review of the literature related to TPES was done in order to determine what the major research themes are for TPES and what the best practices are for it. The first major theme discussed in the literature was that TPES exists for the purpose of increasing teacher performance and student achievement. Not only is this purpose codified into multiple state statutes that are the foundations of statewide evaluation requirements but it is found time and again throughout the literature itself. The second major theme from the literature was the problems and

risks associated with TPES. There is disagreement amongst teachers as to whom the evaluator should be that conducts classroom observations. There also is risk associated with the use of value added measurement (VAM) data. VAM data is not perfect and there are limitations for its use due to the difficulty of isolating the impact that a specific teacher has on the achievement of his or her students. The third and final theme of the literature is the compilation of research-based valued practices that effective teacher performance evaluation systems use.

The next step of this capstone is to evaluate whether or not these six practices are being used in Twin Cities metro area public schools. The next chapter of this project describes how that determination will be made.

CHAPTER THREE

Methods

The literature review identified six research-based valued practices that help to create effective teacher performance evaluation systems (TPES). It also discussed possible problems associated with TPES as well as what researchers have found to be the core purpose and benefits of TPES. This chapter lays out the action research methods that will be used to conduct a survey of local teachers in Twin Cities metro area public schools. It also outlines what the setting is for this research action plan and describes the subjects that will participate in it. The pragmatic research worldview will also be discussed as it directly impacted the development and design of this project. Additionally, this chapter describes why a quantitative research method will be used as well as the benefits of using a survey as the measurement tool.

Last, each question of the survey will be discussed and connected to the identified valued practices. The purpose of creating and analyzing this survey will be to address the research question of this capstone: *Are local Twin Cities metro area public schools using research-based valued practices in their teacher performance evaluation systems?*

Research Settings and Subjects

The subjects who were surveyed are teachers at a Twin Cities metro area public high school. The school is in a suburb and has approximately 2,000 students. The scope of this project is limited to one school as the resources do not exist to

allow for the surveying of teachers in many districts and schools. In working with the administration at the district and school level, I was able to gain permission to send the survey out to all teachers at the specific high school. The scope of the project, its purpose, and the potential benefits of the project were discussed with the administrators and was also outlined in the informed consent letter that was sent to all teachers in the school by email. There were 29 subjects who fully or partially completed the survey. All of the respondents are licensed Minnesota high school teachers who teach various grades and subjects. A breakdown of the grade and subject matter taught by each subject is in Appendix C. Before outlining the survey itself and the specific question, the capstone will first discuss the worldview that drove the creation and design of the survey.

The Pragmatic Worldview

As described by Creswell (2009), researchers are impacted by their own worldviews because worldviews themselves are a “general orientation about the world and the nature of research” (p. 6). This project is derived and carried out from the pragmatic worldview. First, a pragmatic researcher tries to use various research methods and approaches to solving a researching problem (Creswell, 2009). This capstone project has the literature review findings and the survey results as two sources of information. Second, a pragmatic researcher may use quantitative and qualitative data in order to answer the research question (Creswell, 2009). The information that was gathered from the literature review was

qualitative as it was a summary of findings from the review itself. The data that resulted from the survey is quantitative due to the design of the survey. Last, this project is operating within the pragmatic worldview because it conforms to Creswell's (2009) description of pragmatic researchers choosing research methods and techniques that "best meet their needs and purposes" (p.11). One of the reasons a literature review and a corresponding survey were chosen for this project is due to these methods aligning directly with the project goal. Another reason is that these methods work well when dealing with limitations of time and resources. Last, this project is operating in the pragmatic worldview because it allows for the inclusion of other paradigms when it is appropriate and helpful (Creswell, 2009). When all of this is combined together, the pragmatic worldview directly impacted the design of the research questions and the research action plan.

Design and Relevance of the Research Action Plan

The research action plan of this capstone is designed to determine if the six research-based valued practices of effective TPES programs are being used in Twin Cities metro area public schools. To make that determination, the choice was made to take a quantitative approach with a 16-question survey. There are multiple reasons for this. First, a quantitative approach to the research question provides specific answers that can be statistically analyzed so an objective-based conclusion can be made. Second, as Creswell (2009) points out, a quantitative approach is appropriate for situations where there is a need to identify "factors that influence an

outcome,” (p.18). The advantage of using a survey is that it is designed to “provide a quantitative or numeric description of trends, attitudes, or opinions or a population,” (p. 145).

The Survey

The survey was sent out to the school email addresses of 89 teachers on February 02, 2015. The email itself is described in Appendix A and it has a link to the SurveyMonkey website where the survey was built. As is described by Creswell (2009), SurveyMonkey is “an online survey tool” where researchers can “create their own surveys” and receive the results as “descriptive statistics” that can be “downloaded into a spreadsheet or a database for further analysis,” (p.149). The survey was accessible to all respondents who click on the link. The results are then compiled by the website. Once the results are in, the results can be downloaded in an Excel spreadsheet and analyzed. Before opening the survey up for respondents, several test runs were completed to ensure the survey was working properly and all test data was purged from the site.

The data collection method for this survey is cross-sectional as the data represents the views of respondents at a single data point. While respondents had 45 days to complete the survey, it is not a longitudinal study as the data is specific to respondents “at one point in time,” (Creswell, p. 146, 2009). Additionally, the sampling design is single-stage as it is sampling “the people directly” (Creswell, p. 148, 2009) and, unlike with multistage sampling, there will be no sampling from

previously identified clusters from within the population. The respondent sample will also not be stratified to conform to a hypothetical larger population based upon certain characteristics or demographic traits. All analyses will be done based the results from all 29 teachers who responded to the survey.

Table 1 lists the survey questions, what type each question is, what the measurement scale is, and the rationale behind each survey question. Two steps were taken to ensure that no respondent completes the survey twice. First, a completion confirmation statement was provided to respondents at the end of the survey. Second, the survey itself was programmed to only allow one survey to be completed from a given computer based upon its IP address. The completion confirmation statement is described in Appendix B.

Table 1 – TPES Valued Practice Survey

Survey Question	Type of Question	Rationale for Question
Please confirm the email address that your survey link was sent to.	Open Response: Space Provided	This question allows for a completion confirmation email to be sent back to the respondent and it ensures only one survey is completed per respondent.

<p>1) How many years have you been teaching?</p>	<p>Multiple Choice:</p> <ul style="list-style-type: none"> A. 1-3 B. 4-9 C. 10-14 D. 15-24 E. 25+ 	<p>This question allows for the possible analysis that results and the TPES design may differ based upon the length of experience a teacher has.</p>
<p>2) What subject matter do you primarily teach?</p>	<p>Multiple Choice:</p> <ul style="list-style-type: none"> A. English/Literature Arts B. Math C. Science D. Social Studies E. Arts/Music F. Health/Physical Education G. Foreign Language H. Media/Technology I. Elementary 	<p>This question will allow for analyzing whether or not TPES valued practices are being used to evaluate teachers in certain subject matter fields but not in other fields.</p>
<p>3) I feel that our evaluation system is designed to help me improve my teaching skills.</p>	<p>Likert Scale: Strongly Disagree (1) to Strongly Agree (5)</p>	<p>This item will provide data as to whether or not the 1st identified TPES valued practice is being implemented: Effective TPES's have a core purpose of improving teacher and student performance.</p>

<p>4) Which best describes the person(s) that perform classroom observations in your school?</p>	<p>Multiple Choice:</p> <ul style="list-style-type: none"> A. A building administrator B. A peer teacher C. Both A and B D. Other E. We do not have classroom evaluations 	<p>This question provides data as to whether or not TPES valued practice #2 is being implemented: Effective use of classroom observations by peer and administrator evaluators.</p>
<p>5) I feel that the evaluators who conduct our classroom observations are trained to know how to conduct an observation.</p>	<p>Likert Scale: Strongly Disagree (1) to Strongly Agree (5) or Not Applicable</p>	<p>This question is designed to evaluate whether classroom observation evaluators are trained, as is called for by TPES valued practice #3.</p>
<p>6) I feel that evaluators who do our classroom observations have a deep understanding of the curriculum I teach.</p>	<p>Likert Scale: Strongly Disagree (1) to Strongly Agree (5) or Not Applicable</p>	<p>Responses to this question will help to determine if classroom observation evaluators are skilled in the curriculum and instruction, as is called for by TPES valued practice #3.</p>

7) I feel that evaluators who do our classroom observations have a deep understanding of effective classroom instruction.	Likert Scale: Strongly Disagree (1) to Strongly Agree (5) or Not Applicable	Responses to this question will help to determine if classroom observation evaluators are skilled in classroom instruction, as is called for by TPES valued practice #3.
8) Are student test results included as part of your overall evaluation rating?	Multiple Choice: Yes or No	This question is looking to determine whether or not student performance data is included in evaluation results, as is called for by TPES valued practice #4
9) I am aware of what is expected of me and what the criteria and goals are that I am evaluated on.	Likert Scale: Strongly Disagree (1) to Strongly Agree (5)	This question is determining whether or not TPES valued practice #5 is being implemented: Clear expectations exist for teachers within the evaluation that are based upon agreed upon criteria and goals.
10) I agree with the goals and criteria that I am evaluated on.	Likert Scale: Strongly Disagree (1) to Strongly Agree (5) or Not Applicable	Responses to this question will help to determine if teachers agree with what they are evaluated on, as is called for by TPES valued practice #5.
11) My evaluation provides me feedback that I can use to improve my teaching skills.	Likert Scale: Strongly Disagree (1) to Strongly Agree (5) or Not Applicable	This question is looking into whether or not the evaluation is giving productive feedback that allows teachers the ability to improve their performance, as is called for by TPES valued practice #6.

12) I receive feedback immediately after an observation.	Likert Scale: Strongly Disagree (1) to Strongly Agree (5) or Not Applicable	The question is determining whether or not the TPES valued practice #6 of having timely feedback is occurring.
13) I receive feedback about my teaching multiple times throughout the entire school year.	Likert Scale: Strongly Disagree (1) to Strongly Agree (5)	This question is looking into whether or not the frequent feedback component of TPES valued practice #6 is occurring.
14) I can seek out professional development opportunities based upon my evaluation feedback.	Likert Scale: Strongly Disagree (1) to Strongly Agree (5) or Not Applicable	This question is determining whether or not their evaluations provide feedback that can be linked to targeted professional development, as is called for by TPES valued practice #6.
15) How are evaluation results delivered to you?	Open Response: Space Provided	This question is looking into what type of setting and method is most frequently being used for communicating results to teachers. This question is creating the potential for additional data above and beyond the valued practices.
16) Is there a change you would like to see occur with your evaluation system?	Open Response: Space Provided	This question is allowing respondents to offer feedback for areas where they may believe there is room for improvement for their evaluation system and this could provide potential new data to be used as a basis for further research.

Several steps occurred following the completion of the survey by respondents. First, the survey on SurveyMonkey was closed to ensure no further participation could occur that would alter the results. Second, the survey data was downloaded into Microsoft Excel spreadsheets where it was compiled and analyzed. Finally, multiple charts and graphs were made in Microsoft Excel and transferred into this document based upon the analyses that were performed.

Summary

The research action plan described in this chapter was designed to answer the following research question: *Are local Twin Cities metro area public schools using research-based valued practices in their teacher performance evaluation systems?* The research plan arose from the pragmatist worldview as it combined multiple sources of data and it is designed based upon specific resource limitations. The research action plan was quantitative in nature because this route allowed for a statistical analysis that can make an objective conclusion regarding the research question. A survey was used as the quantitative instrument because it allowed respondents to give their opinion as to how their TPES operates in relation to the six TPES research-based valued practices that were described in the literature review. Results and analysis from the 16-question survey will be discussed in the next chapter.

CHAPTER FOUR

Research Results

This chapter will discuss the results of the survey and what the data communicates regarding the use of research-based valued practices for teacher evaluations within Twin Cities metro area public schools. It will first describe how the survey completion process occurred, the amount of data that was collected, and the specific responses. Then it will analyze and interpret how the results apply to the research question by defining trends and patterns that exist within the data.

Survey Collection and Results

After gaining district level and school administrator permission, the “survey request email” (informed consent) was sent to all staff within a suburban 9-12 high school that has approximately 2,000 students and 89 teachers. The survey was open for 45 days and 29 teachers completed it.

Table 2 shows the results from the survey questions. Not all questions were answered by the same number of respondents, only two questions required a response, and the response percentage does not always add up to 100% due to rounding. The open-ended responses for the final two survey questions are compiled in Appendix D.

Table 2 - Survey Results

Survey Question	Possible Response	Response Count	Response %	Response Average
1) How many years have you been teaching? (29 responses)	1-3 years	2	6.9%	N/A
	4-9 years	10	34.5%	
	10-14 years	9	31.0%	
	15-24 years	6	20.7%	
	25+ years	2	6.9%	
2) What subject matter do you primarily teach? (27 responses)	English/Literature Arts	7	25.0%	N/A
	Math	6	21.4%	
	Science	4	14.3%	
	Social Studies	1	3.6%	
	Health/Physical Ed.	1	3.6%	
	Foreign Language	1	3.6%	
	Media Technology	1	3.6%	
	Elementary	0	0%	
	Special Education	6	21.4%	
3) I feel that our evaluation system is designed to help me improve my teaching skills. (27 responses)	1 – Strongly Disagree	1	3.7%	3.44
	2 – Disagree	4	14.8%	
	3 – Neutral	6	22.2%	
	4 – Agree	14	51.9%	
	5 – Strongly Agree	2	7.4%	
4) Which best describes the person(s) that completes classroom observations in your school? (27 responses)	A building administrator	7	25.9%	N/A
	A peer teacher	3	11.1%	
	A building administrator and a peer teacher	17	63.0%	
	Other	0	0.0%	
	We do not have classroom evaluations	0	0.0%	
5) I feel that the evaluators who conduct our classroom	1 – Strongly Disagree	1	3.7%	3.78

observations are trained to know how to conduct an observation. (27 responses)	2 – Disagree	1	3.7%	
	3 – Neutral	5	18.5%	
	4 – Agree	16	59.3%	
	5 – Strongly Agree	4	14.8%	
6) I feel that evaluators who do our classroom observations have a deep understanding of the curriculum I teach. (27 responses)	1 – Strongly Disagree	4	14.8%	2.52
	2 – Disagree	12	44.4%	
	3 – Neutral	4	14.8%	
	4 – Agree	7	25.9%	
	5 – Strongly Agree	0	0.0%	
7) I feel that evaluators who do our classroom observations have a deep understanding of effective classroom instruction. (27 responses)	1 – Strongly Disagree	1	3.7%	3.70
	2 – Disagree	3	11.1%	
	3 – Neutral	4	14.8%	
	4 – Agree	14	51.9%	
	5 – Strongly Agree	5	18.5%	
8) Are student test results included as part of your overall evaluation rating? (27 responses)	Yes	2	7.4%	N/A
	No	25	92.6%	
9) I am aware of what is expected of me and what the criteria and goals are that I am evaluated on. (27 responses)	1 – Strongly Disagree	0	0.0%	4.19
	2 – Disagree	2	7.4%	
	3 – Neutral	2	7.4%	
	4 – Agree	12	44.4%	
	5 – Strongly Agree	11	40.8%	
10) I agree with the goals and criteria that I am evaluated on. (27 responses)	1 – Strongly Disagree	0	0.0%	3.78
	2 – Disagree	1	3.7%	
	3 – Neutral	8	29.6%	
	4 – Agree	14	51.9%	
	5 – Strongly Agree	4	14.8%	
11) My evaluation provides me feedback that I can use to improve my teaching	1 – Strongly Disagree	0	0.0%	3.41
	2 – Disagree	5	18.5%	

skills. (27 responses)	3 – Neutral	8	29.6%	
	4 – Agree	12	44.4%	
	5 – Strongly Agree	2	7.4%	
12) I receive feedback immediately after an observation. (27 responses)	1 – Strongly Disagree	0	0.0%	3.30
	2 – Disagree	10	37.0%	
	3 – Neutral	3	11.1%	
	4 – Agree	10	37.0%	
	5 – Strongly Agree	4	14.8%	
13) I receive feedback about my teaching multiple times throughout the entire school year. (26 responses)	1 – Strongly Disagree	1	3.8%	3.08
	2 – Disagree	10	38.5%	
	3 – Neutral	3	11.5%	
	4 – Agree	10	38.5%	
	5 – Strongly Agree	2	7.7%	
14) I can seek out professional development opportunities based upon my evaluation feedback. (26 responses)	1 – Strongly Disagree	0	0.0%	3.69
	2 – Disagree	2	7.7%	
	3 – Neutral	7	26.9%	
	4 – Agree	14	53.8%	
	5 – Strongly Agree	3	11.5%	
15) How are evaluation results delivered to you? (26 responses)	See Appendix D for open ended responses			N/A
16) Is there a change you would like to see occur to your evaluation system? (22 responses)	No	8	36.4%	N/A
	See Appendix D for open ended responses			

The Evaluators

Several trends within this dataset are apparent when analyzing them in conjunction with the overall survey design and the research-based valued practice

that was being asked about with each question. The first trend relates to how the respondents felt about the classroom evaluator who conducts the evaluations. When respondents were asked whether they felt the evaluator “had a deep understanding of the curriculum” being taught, the response average was 2.52. This is below the “neutral” rating of 3.0. Conversely, the average ratings for whether or not they felt evaluators have “a deep understanding of effective classroom instruction” (3.70) and were “trained to know how to conduct an evaluation” (3.78) were both above the “neutral” rating. The data also shows that most respondents feel that evaluations are done by both a building administrator and a peer teacher (63%). Overall, in three of the four areas that involve the traits of the classroom observers, the data shows that research-based valued practices are being used within this school.

The Evaluation

The second trend involves how respondents felt about the design of their evaluations. When asked if they felt their evaluation was designed to help their teaching skills, the average rating 3.44 and 16 of the 27 respondents said they “agree” or “strongly agree.” Regarding whether they felt they were “aware of what is expected” of them and “the criteria and goals” they are evaluated on, the response average was 4.19 with 23 of 27 respondents saying they “agree” or “strongly agree.” As to whether respondents agreed with the “goals and criteria” they are evaluated on, the response average was 3.78 and 18 of 27 respondents said they “agree” or “strongly agree.” The data shows continued alignment with TPES valued practices as

teachers within this school are very aware of what they are evaluated on, they agree with what they are evaluated on, and they overall believe the design of the evaluation is meant to help them improve their teaching skills.

Valuable Feedback

The third major trend involves how respondents felt about the flow of evaluation feedback and how it may allow them to improve their own teaching skills. With 14 of 27 respondents saying “agree” or “strongly agree” and an overall response rating of 3.30, the surveyed teachers overall felt they did “receive feedback immediately after an observation.” The data also shows a response rating of 3.41 when respondents were asked if the evaluation feedback they are given can be used to “improve (their) teaching skills.” However, while the valued practice is to have this feedback occur multiple times throughout the school year, respondents were nearly closely divided as to whether they agreed or disagreed that this was happening with their own evaluations.

Eleven of 26 respondents said they “disagree” or “strongly disagree” while twelve respondents said they “agree” or “strongly agree” that they received feedback multiple times during the school year. With three “neutral” responses, this question had a rating average of 3.08. It is not clear from this data how or why respondents would be so clearly divided on this topic. Possible reasons for the divide could be disagreement by respondents as to what “multiple” means or perhaps that frequency of feedback does indeed vary from teacher to teacher within

this school. Last, when asked if they felt they could “seek out professional development opportunities based upon” their evaluation feedback, 17 of 26 respondents said they “agree” or “strongly agree” that they could and the response average of 3.69. Overall the data shows that the evaluation system in place within this school is connecting feedback to improvement opportunities but the frequency and timing of the feedback does not appear to completely align with the valued practice of being immediate and frequent.

The last trend that is apparent from the survey data involves the use of student data within the evaluation itself. When asked if student test results are a part of their “overall evaluation rating,” 25 of 27 respondents said “no.” This does not align with the research-based valued practice of including limited but valid student test data within the evaluation that teachers receive. When asked if there is a change he/she would like to see to the evaluation system, respondent #19 stated, “Do not tie it into student results. Student results have mostly to do with their abilities, backgrounds, and the teacher should not be evaluated on that.” While no other open-ended comments directly spoke to the use of student data on teacher evaluations, it is clear from the data that it is not being done within this school.

Relationships within the Data

While the analysis above shows that the evaluation system in place within the surveyed school overall appears to align with research-based valued practices,

this discussion will now look into the traits of who answered in certain ways and if certain relationships exist within the data.

One of the data points that had the most striking contrast with responses was whether respondents agreed or disagreed that evaluation feedback was provided “immediately after an observation.” Ten respondents disagreed with their feedback being immediate, ten agreed, and four strongly agreed it was immediate. In looking deeper at the data, there does not appear to be a significant correlation between those responses and the experience level or subject area of the respondents. This is evident by the tables below:

Table 3 – Immediate Feedback & Experience

“I receive feedback immediately after an observation”		
Years of Experience	# of “Agree/Strongly Agree” Responses	# of “Disagree/Strongly Disagree” Responses
1-3 Years	0	1
4-9 Years	4	4
10-14 Years	5	3
15-24 Years	4	1
25+ Years	1	1

Table 4 – Immediate Feedback and Subject Matter

“I receive feedback immediately after an observation”		
Primary Subject Matter	# of “Agree/Strongly Agree” Responses	# of “Disagree/Strongly Disagree” Responses
Math	2	3
Science	1	3
English/Literature Arts	4	2
Special Education	3	1
Social Studies	0	1
Media/Technology	1	0
Arts/Music	1	0
Health/Physical Education	1	0
Foreign Language	1	0

While English/Literature Arts teachers were twice as likely to “Agree/Strongly Agree” that their feedback was immediate versus “Disagree/Strongly Disagree” and Special Education teachers were three times as likely to do the same, the number of respondents for both subject matters is too small to make overall conclusions related to this relationship.

Another survey question that has a striking gap with respondent answers is whether teachers agreed or disagreed that they receive feedback multiple times per year. As was described previously, 11 respondents “Disagree/Strongly Disagree” with the statement of “I receive feedback about my teaching multiple times throughout the entire school year” whereas 12 respondents “Agree/Strongly Agree” with that statement. A breakdown of these responses by experience level and by primary subject matter is shown in the tables below:

Table 5 – Multiple Feedback and Experience

“I receive feedback about my teaching multiple times throughout the entire school year”		
Years of Experience	# of “Agree/Strongly Agree” Responses	# of “Disagree/Strongly Disagree” Responses
1-3 Years	0	1
4-9 Years	7	2
10-14 Years	3	4
15-24 Years	2	2
25+ Years	0	2

Table 6 – Multiple Feedback and Subject Matter

“I receive feedback about my teaching multiple times throughout the entire school year”		
Primary Subject Matter	# of “Agree/Strongly Agree” Responses	# of “Disagree/Strongly Disagree” Responses
Math	3	1
Science	1	3
English/Literature Arts	3	2
Special Education	2	2
Social Studies	0	1
Media/Technology	1	0
Arts/Music	1	0
Health/Physical Education	0	1
Foreign Language	0	1

While there does not appear to be a significant correlation between these responses and the primary subject matter the respondents are teaching, there does appear to be a trend regarding years of experience. When analyzing the 10 responses from teachers who have 1-9 years of experience, 70% stated they “Agree/Strongly Agree” that they receive feedback multiple times during the school year. Conversely, for the 13 responses from teachers with 10+ years of experience, only 38% stated “Agree/Strongly Agree” that feedback occurs multiple times during the school year. There clearly is a trend that more veteran teachers are less likely to receive feedback multiple times in a school year. A possible explanation that could be researched further is whether this trend is due to contractual obligations.

A final area where that data presents to us an interesting relationship involves how respondents felt about the evaluators who conduct the classroom observations. It was previously discussed that there are two questions involving the

evaluators that have a wide margin in the data. The two questions are: “I feel that evaluators who do our classroom observations have a deep understanding of the curriculum I teach” (response average of 2.52); and “I feel that evaluators who do our classroom observations have a deep understanding of effective classroom instruction,” (response average of 3.70). Figures 1 and 2 show the distribution of responses for these two questions by the primary subject being taught for each respondent and when neutral responses are not included. Subject areas with less than two responses are not included in these figures.

Figure 1 - The evaluator has a deep understanding of effective classroom instruction

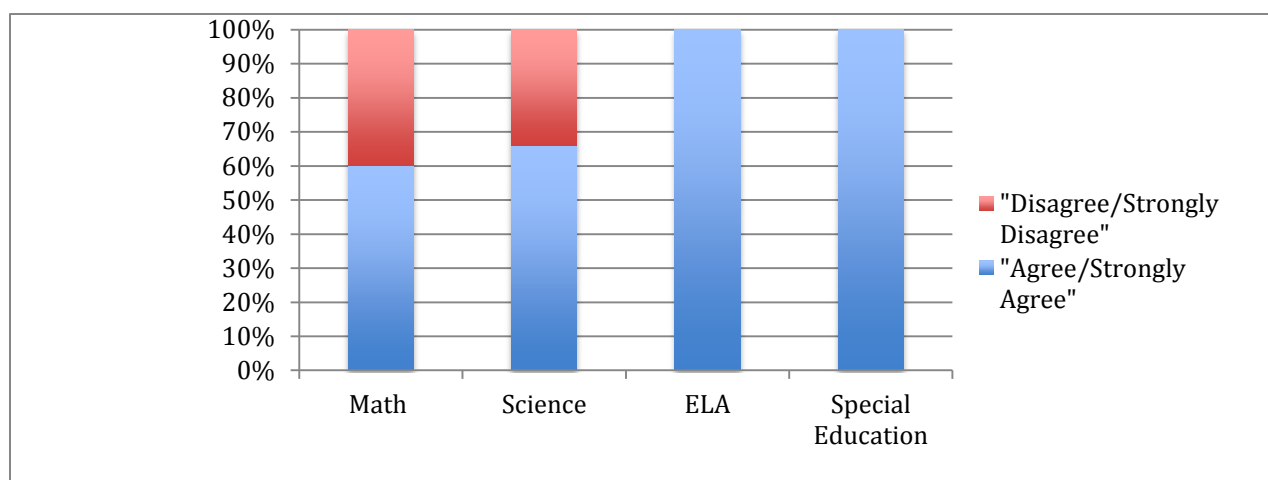
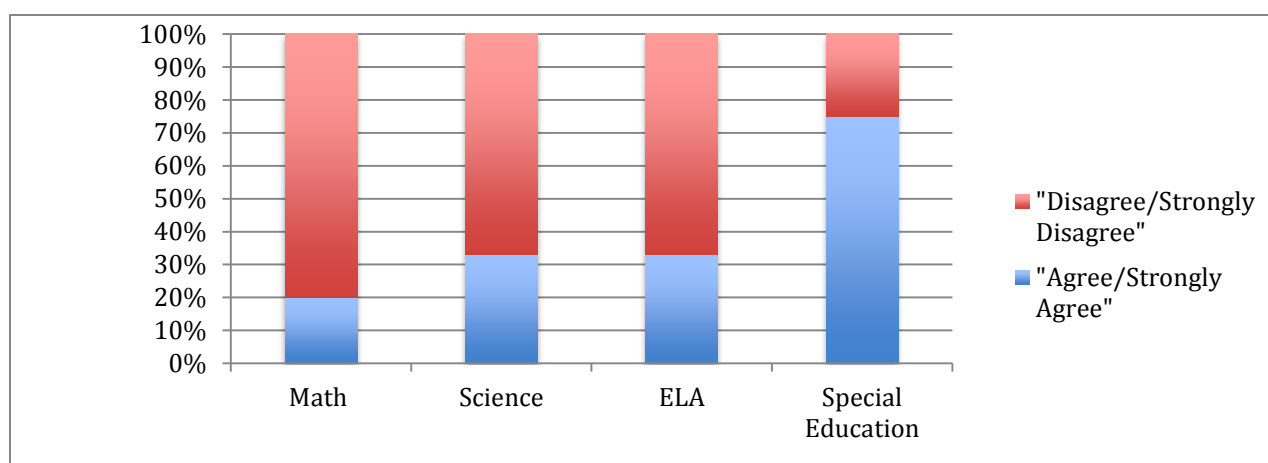


Figure 2 - The evaluator has a deep understanding of the curriculum I teach



Figures 3 and 4 lay out the data in the same manner for the same two questions, the neutral responses are not included, but it is distributed by the years of experience of the respondents. Experience categories with one or two responses are not included .

Figure 3 - The evaluator has a deep understanding of effective classroom instruction.

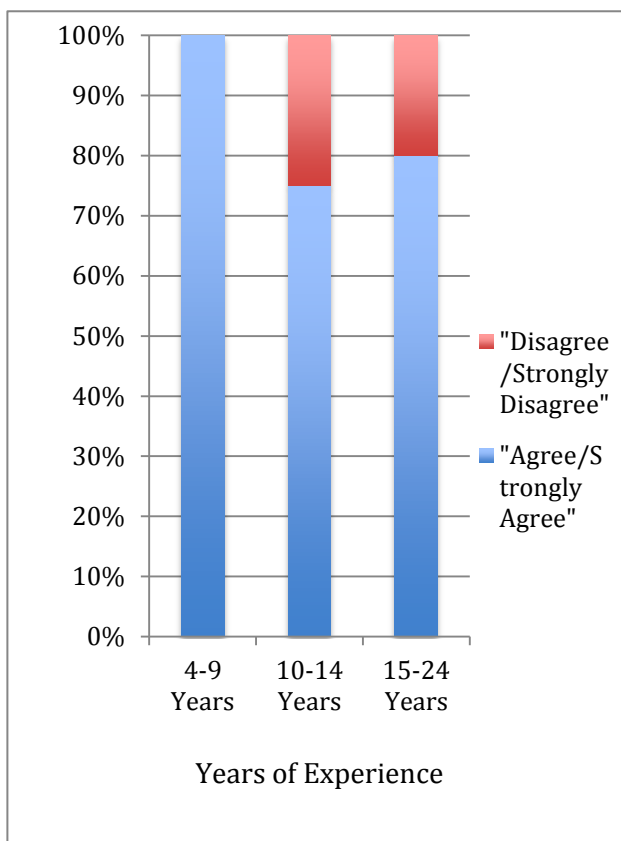
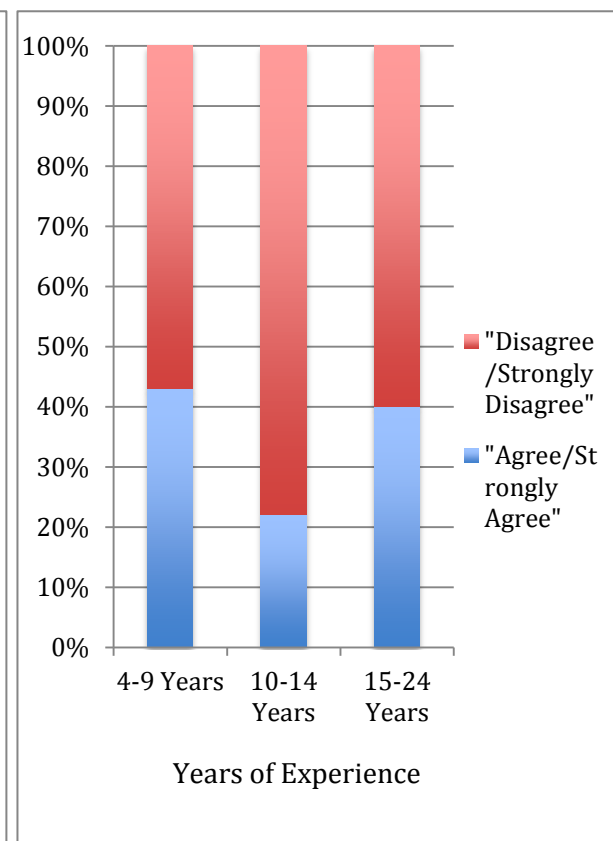


Figure 4 - The evaluator has a deep understanding of the curriculum I teach.



These four figures indicate that regardless of the years of experience or the subject matter the respondent teaches, there is more confidence in the evaluators'

understanding of effective classroom instruction than there is in their understanding of the curriculum being taught.

One possible reason for this data could be that evaluators do not evaluate teachers in classrooms that they themselves are content specialists in. This would eliminate the chance of the evaluator possibly knowing the teacher and that relationship could impact the observation. Another possible reason may be that evaluators pick classrooms completely at random, regardless of the experience level or subject matter being taught. This random sampling of classrooms for evaluation would ensure that the process and evaluation data are statistically valid. Further, as there are many possible subject areas and evaluators likely only specialize in the one content area that they personally have experience with, a random assigning of evaluators would mean that most evaluators are evaluating a teacher who is teaching a subject area that the evaluator is not an expert in.

Results Conclusion

Several trends and relationships can be noted as existing within the sampled school based upon the survey data. First and foremost, research-based TPES valued practices appear to overall be in place within the school. Teachers felt their evaluators were trained in how to complete evaluations and that they understood effective classroom instruction but teachers did not feel that observers knew the subject matter being taught. Teachers overall agreed with what they were evaluated on and were aware of what criteria and goals they were evaluated on but student

data does appear to be used within this TPES. Teachers also felt that their evaluation was meant to help them and that it connected with professional development even though nearly a majority felt they did not receive feedback immediately after an observation or multiple times per year.

Teachers with less than ten years of experience were nearly twice as likely to feel their feedback occurred multiple times per year but were just as split as the other experience categories when asked if they felt their feedback was immediate. Additionally, given how high the response rate was within the sampled school, the potential does exist that if the survey were able to expand to other schools, there could be a very large and valuable data set to use for many more analyses and for determining larger trends. While the next chapter will discuss some of the limitations of this data as well as recommendations for future research, the data does help to show this school overall does have a TPES that aligns with research-based valued practices.

CHAPTER FIVE

Conclusion

The goal of this capstone project was to answer as best as possible the following research question: *Are local Twin Cities metro area public schools using research-based valued practices in their teacher performance evaluation systems?* After an in-depth review of the pertinent literature and outlining the design of the survey project, this capstone described that, based upon the data that was collected, the answer to this question is yes. Nearly all of the valued practices identified in the literature review for an effective TPES were found to be present within the sampled school. Before sending out the survey and getting these results, the first step was to review the pertinent literature to synthesize what the research-based valued practices are.

Literature Review

The second chapter of this capstone discussed the pertinent information that currently exists within the literature for the topic of teacher performance evaluation systems. The literature review substantiated that there are several common themes, including what this project calls “valued practices,” within the literature when it comes to TPES.

The first major theme that the literature discussed was that effective TPES exists in an effort to increase teacher performance and student achievement. This theme was found in the literature itself and within the statutes of multiple states. The second major theme involved the problems and risks that are associated with

TPES. The main factors here involve the risk of using value-added measurement data (VAM), disagreement about who the evaluator should be that conducts classroom evaluations, and how a teacher's evaluation should impact their employment. The final theme that was found within the literature is the compilation of six valued practices for effective TPES within a school.

The six valued practices that help to create an effective TPES per the literature are:

1. to have the purpose of increasing teacher and student performance through evaluation feedback (Olivia et al, 2009; Weisberg et al, 2009);
2. the effective use of classroom observations by peer and administrative evaluators (Chukwubikem, 2012; Conley, 1987; Hiller, 1986; Kane et al., 2011; Mathers & Olivia, 2008);
3. having trained and qualified evaluators conduct classroom observations (Conley, 1987; Mathers & Olivia, 2008; Olivia, Mathers, & Lane, 2009; Stark & Lowther, 1984; Weisberg et al., 2009; Zakariya, 1985);
4. proper and specific use of VAM data (Kersting, Chen, & Stigler, 2013; Maslow & Kelley, 2012; Mathers & Olivia, 2008);
5. having clear expectations of agreed upon criteria and goals within the evaluation (Chukwubikem, 2012; Looney, 2011; Stark & Lowther, 1984; Weisberg et al., 2009); and
6. providing timely and frequent evaluation feedback that is linked directly to professional development to allow for targeted teacher growth (Benedict et

al., 2013; Chukwubikem, 2012; Delvaux et al., 2013; Looney, 2011; Mathers & Olivia, 2008; Stark & Lowther, 1984; Weisberg et al, 2009).

By identifying these six common valued practices within the literature, specific questions were then designed and built into the survey so that the research question could be answered. The survey showed that a majority of the valued practices and the subcomponents are being used within the sampled school. The only practices that the data does not show support for being used within the sampled school involve the proper and specific use of VAM data and making sure that the evaluators are qualified in having a deep understanding of the subject matter being taught.

Implications of Study

Based upon the research, survey, data collection, and analyses that were done within this capstone, several implications arise. First, there is a possibility that the sampled school is not unique within the Twin Cities metro area and that many more public schools are using the identified research-based valued practices for effective TPES. As all other public schools in the Twin Cities (and in Minnesota overall) are subject to the same state statute that mandates use of TPES, it is not unreasonable to consider the possibility that the sampled school is similar to other public schools in the Twin Cities. However, in order to validate this possibility, additional research at a larger scale would need to be conducted.

Second, based upon the findings of this capstone, there is a possibility that

implementation of many of the valued practices is a product of factors that are unique to this specific school and not to others. It is not known based upon the survey what factors may be present within this school to have caused the evaluators to have a deep understanding of classroom instruction, or what caused the teachers to know what they are evaluated upon and to agree with it, or what caused the evaluators to consist of a peer and an administrator. These practices may be based upon factors that are unique only to this school, such as through the choices of specific school administrators or teachers. Determining what the driving forces are for the usage of these practices is another area for further research.

Another implication of this study is the possibility that TPES is a concept that is being paid attention to by teachers within the Twin Cities area. This is a possibility based upon the survey completion rate. Of the 89 teachers that were sent the email that described the project overall and provided the survey link, 29 teachers completed it. This is a completion rate of 32.6%. The expectation going into the project was that the completion rate within a sampled school would be 10-20%. Because there was no benefit to respondents to complete the survey, there exists the possibility that an outside factor impacted the completion rate. One such possibility for the completion rate is that TPES is very much on the minds of teachers within the sampled school and that these teachers felt a need to communicate their opinions about their own TPES and what their experiences have been.

Limitations of Study

As with any research project, there are limitations that need to be acknowledged regarding the analysis of this data. The first limitation is that the survey was completed within one school. While the number of respondents and the completion percentage within the school is acceptable, the sampling was done within one school in one school district. Efforts were made to increase the number of participating schools but time and resource constraints required that the project move forward with the single school. The impact of the data arising from one school is that while valuable perspectives and information can be gained regarding the application of research-based valued practices for TPES within this location, it is not automatically presumed to be like this in other Twin Cities public schools. This school is similar to many other large suburban schools in terms of number of students, the racial makeup of the student body, the salary scales for teachers, and the demographics of the community overall. But, it cannot be presumed that those similarities would translate into similar TPES practices being used in those other locations. An area for additional study would be to replicate this survey and expand it to other schools so larger trends can be determined.

Another limitation of this study involves the sample sizes for the subject area and years of teaching categories. The purpose of asking respondents for these two traits was to see if survey answers correlated with the subject area and/or experience level of respondents. If the survey had been completed within multiple schools, there likely would not be categories with only one or two respondents.

While categories like “4-9 years” (ten), “10-14 years” (nine), and “English/Literature Arts” (seven) had high numbers of respondents, categories such as “25+ years” (two), “social studies” (one) and “health/physical education” (one) did not. Because there was a variation in the number of respondents in the sampled categories, Figures 1, 2, 3, and 4 did not include categories with two or fewer respondents. Not including categories with two or fewer respondents within those figures means the analysis was limited. If this survey were to be repeated with a larger set of schools, another potential benefit would be the ability to analyze with more subject areas and more experience level categories.

Another potential limitation of the study involves the chance that personal biases from myself as the researcher may have impacted the overall outcome. As was described earlier in the capstone, I currently work in performance evaluation, I previously was a teacher, and I was then and remain today strongly in favor of having teachers evaluated for the benefit of improving their skills as well as student performance. However, through proper research methodology and a rigorous peer review process, any potential bias that may impact the capstone outcome has hopefully been negated.

Lastly, there is the possibility that the data is skewed due to intrinsic motivations that respondents may have regarding TPES. While a survey completion rate of 32.6% was much higher than was expected, it still means that over 2/3 of teachers did not complete the survey. The possibility exists that the teachers who did not complete the survey feel very different about the school’s TPES than those

who did complete it. The possibility also exists that if the views of the other 67.4% had been sampled, the data would have been even stronger in terms of showing that the research-based valued practices for effective TPES are being used.

Future Research

The limitations of this study directly connect with areas that could be researched in the future. In order to help understand the situation at a larger level, future research could include expanding this survey to more school districts and to more schools. Expanding the sample size would allow for additional analyses of the data. One additional analysis involves looking deeper at the responses of teachers when matched against identified traits. In this survey, years of experience and subject area were sampled on the possibility that they may correlate to certain responses. Having a larger sample with more schools would eliminate this problem and allow for those analyses to be made for all sub-groups.

If more schools were sampled, another area for additional research that was described previously involves determining why, or why not, certain practices are being implemented within a specific school or district. Once many more schools were sampled, the first step would be to analyze the data and determine what differences exist between the schools in terms of what practices are being implemented and which ones are not. Based upon those differences, additional research would occur to determine if there are factors within the schools that would cause that difference. One way of doing this would be to interview administrators

and teachers at each school and determine what the factors are that led to the current state of their TPES. Common factors may be found that, when present, lead to certain valued practices being implemented no matter which school it is. Conversely, there may be no factors that exist at the school level that lead to certain practices being implemented.

A final area for additional research that arises based upon this project involves surveying school administrators instead of teachers. This project surveyed teachers as they are the ones being evaluated within a TPES. However, administrators are involved with a TPES and have a direct role in its creation and maintenance. Not only would it be interesting to determine if there are differences in the responses from teachers and school administrators, but documenting those possible differences and even looking into why they exist would allow for everyone involved in a TPES to have a better understanding of what is going on, where people disagree, and allow for improvements and updates to be made.

Personal Growth & Learning

This capstone has been not only an opportunity to add to the literature but also a chance for myself to grow as a researcher. People may have prior experiences that can aid and assist in making the capstone process easier than it is for others. However, even with past experiences aiding in the capstone process, I have found this entire process to be a learning and growth opportunity.

One area of growth involved making the text as plain language as possible. It

cannot be presumed that a grammatically correct sentence is also the best way of communicating to a reader. Through the review and revision process with my committee, I have been able to see how my own writing can be simplified and shortened while still conveying the same message. The simplifying and shortening of my writing is somewhat new to me even as a professional who creates reports and papers on a frequent basis. I have found in certain situations there is a demand for writing to be drawn out and lengthy in order to ensure every possible point is conveyed to the reader and to decrease the chance of information being missed. For this capstone, I have learned that more is not better and there is not necessarily a need to reinforce a point several times.

I was also able to grow as a thinker through the capstone process. Receiving consistent and quality feedback throughout the process from committee members allowed me ample opportunity to rethink nearly the entire capstone from beginning to end. Each time I submitted a new portion or a new chapter for review, I thought it was exactly how I wanted it and there would not be a need for significant changes. The capstone process has taught me to think differently of my research when developing it and then writing about it. Instead of writing about it in terms of how I want to describe it, I have learned to also consider what my readers would want to read as well. This relates to the concept of putting yourself in the chair of your intended audience. I as the researcher know a vast amount of information regarding TPES but I cannot presume that every person reading this capstone knows that

same information. Realizing this difference of perspective is essential to writing a quality capstone that can be understood by a reader.

A final area of growth for myself through this process involves the research process. I was very familiar with the scientific method, statistical analysis, and social research before starting this project. But, having a deep understanding of how to create the research project itself, how to carry out the research, and how to document it are things in my view that can only be learned by actually doing them. I can read from many different books how to create a research project but actually doing it is a different concept. Reading about the research process on paper can be very abstract and I feel that one needs to actually do it to truly understand the many facets of effective and valid social research.

Communication of Results

The survey data and the overall findings from this capstone will be made available to several groups. The capstone will be made public within the Hamline University library system so any future students, researchers, faculty, or members of the public can look into what this capstone added to the literature. The data will also be made available to specific groups related to the sampled school. As per the research agreement that was made with the school district's administration, the data will be sent to them and no reference will be made within this capstone about the specific identity of the school that was sampled. The administration at the sampled school will also have the data made available to them as they requested it

to help in their own efforts to update and reform their TPES. By providing access to this capstone in many ways to many different users, hopefully it can be used to help expand the use of research-based TPES valued practices. This continued expansion may provide assurances to interested stakeholders regarding teacher performance quality.

Conclusion

The goal of this project was to determine if research-valued practices for effective teacher performance evaluation systems are being used in Twin Cities metro area public schools. The literature review provided six commonly used effective practices that allowed for the development of the survey. Data from the survey showed that overall the valued practices are indeed being used at the school that was surveyed. While there are some limitations to this study regarding its applicability to the overall metro area, this study can be used as a starting point for additional research that can delve into this topic further. At the end of the day, there are stakeholders who want to know how well American public school teachers are performing. Effective teacher performance evaluation systems are one way that school districts can report how its teachers are doing beyond student test scores. So far, it appears that effective practices are being used to help tell this story to the American society.

Appendix A - Survey Request Email (Letter of Informed Consent)

Mr./Ms. XXX,

My name is Jeff Holtz and I am a Master's degree seeking student in the Masters of Arts in Teaching program at Hamline University. I am conducting research into teacher evaluation systems that are being used in Twin Cities metro area public schools and whether research-based valued practices are being applied.

XXX indicated that it would be permissible to ask staff to complete a 16-question survey about your teacher evaluation system. The survey should take less than 15 minutes to complete and there is no risk to you as a participant of this survey.

Your responses are confidential, they will be used for research analysis purposes only, and all names will be changed to numeric designations. This research is public scholarship and will be cataloged in the Bush Library Digital Commons at Hamline University. If you wish to participate in this research project, please click on the following link and answer the 16 questions: *(link)*.

You can withdraw from the survey at any point prior to completing the survey. Upon completion you may print off your survey responses for future reference.

Thank you.

Jeff Holtz

[Redacted]
[Redacted]
[Redacted]

Saint Paul, MN 55119

M.A.T. Candidate, Hamline University

(Consent statement from the school administration)

Appendix B - Survey Completion Confirmation Statement

Your survey is now complete. As a reminder, your responses are confidential and will be used for research analysis purposes only. Thank you for your participation.

Jeff Holtz

M.A.T. Candidate, Hamline University

Appendix C – Breakdown of Respondents by Experience & Subject Matter

Years of Experience	Primary Subject Matter
10-14 years	Arts/Music
4-9 years	English/Literature Arts
15-24 years	Math
1-3 years	Science
4-9 years	Special Education
10-14 years	English/Literature Arts
4-9 years	Math
4-9 years	Special Education
10-14 years	Science
10-14 years	Media/Technology
1-3 years	English/Literature Arts
25+ years	Special Education
4-9 years	English/Literature Arts
4-9 years	Math
15-24 years	Foreign Language
15-24 years	English/Literature Arts
10-14 years	Science
15-24 years	<i>(no response)</i>
10-14 years	Math
10-14 years	Health/Physical Education
10-14 years	Special Education
15-24 years	Special Education
15-24 years	English/Literature Arts
25+ years	Math
10-14 years	Social Studies
4-9 years	Math
4-9 years	English/Literature Arts
4-9 years	Science
4-9 years	Special Education

Appendix D – Open Ended Responses

How are evaluation results delivered to you?	Is there a chance you would like to see occur to your evaluation system?
Post evaluation conference	I would like to be evaluated by professionals in my subject area.
In person meetings	No
In person.	The current evaluation process is for administrative purposes.
One on one conference and a copy of the results in mailbox	More consistency, less paperwork, more immediate feedback, more informal walk throughs and feedback, better understanding of curriculum concepts and goals
meeting	No
When I am fully evaluated, I am given the results in a meeting with the administrator who completed my evaluation. When I am peer-evaluated (informally), I arrange with that peer to go over the results.	Yes. I would like to see more changes in who is evaluating each teacher. I have been fortunate enough to be evaluated by someone I feel to be capable and competent in that role, but there are those among my colleagues who are stuck being evaluated by those who are not, and there doesn't seem to be much of an opportunity to be evaluated by someone new...it's unfair for us to be denied fair access to capable evaluators.
If a formal observation with an administrator, you typically have a scheduled meeting with the administrator in which you discuss the observation. If it is a peer observation, the results may only be delivered via a form in your mailbox.	It would be nice if a content specialist observed the appropriate content area.
In person	Be evaluated more than every 3 years and be given specifics I can work on, not that I'm good at it.
at a face to face meeting several weeks after the fact.	Have it based on what I do, and immediate
email, mailbox, paper, in-person	In-District Peer Evaluation; so I could have a teacher who deeply understands my curriculum and could add valuable feedback to help me improve my teaching and the experience for my students.
<i>(no response)</i>	
in person	Yes, I think we are evaluated too much. If you are a proven teacher then there should be more focus on

	the young teachers.
Verbal feedback, completed forms and rubrics (hard copy, not online)	Not that I can think of now. I don't think the system is entirely effective, but I can't think of a better way to do things.
With formal administrative observations, there is a post-observation meeting where we go through the results. For peer observations, the results sheet is normally given to us before the observer leaves the room or the next day.	I would like the administrative observations to last more than 5-10 minutes, even for the "walk-throughs", which are supposed to be less formal and shorter. I say this because it's hard to give good advice when you're only observing a small portion of the class period.
conference with papers	No.
face to face meeting, electronic forms	No
face to face meeting	no
<i>(no response)</i>	<i>(no response)</i>
In a person discussion with my peers and my supervisor	No. It's been greatly improved over the last few years.
face to face meeting	<i>(no response)</i>
<i>(no response)</i>	<i>(no response)</i>
Face to face and print	No
Meetings.	<i>(no response)</i>
in person	yes
paper mail	Do NOT tie it into student results. Student results have mostly to do with their abilities, backgrounds, and the teacher should not be evaluated on that.
Conference format for discussion and also recorded in our online PD record keeping system.	<i>(no response)</i>
In writing.	<i>(no response)</i>
Usually in a one-on-one meeting setting	I feel that teacher to teacher observations could be changed to be more efficient and effective.
In person	No

References

- Benedict, A.E., Thomas, R., Kimerling, J., Leko, C. (2013). Trends in teacher evaluation: what every special education teacher should know. *TEACHING Exceptional Children*, 45(5), 60-68.
- Brandt, C., Mathers, C., Olivia, M., Brown-Sims, M., & Hess, L. (2007). Examining district guidance to schools on teacher evaluation policies in the Midwest Region (Vol. 30). Washington D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest.
- Byrant Research LLC. (2013). *2013 education poll*. Retrieved from http://www.knoxvillechamber.com/sites/default/files/randys_presentation.pdf
- Bushaw, W. J., & Lopez, S. J. (2012). Public education in the United States: A nation divided. *Phi Delta Kappan*, 94(1), 8-25.
- CBS News Corporation. (2010). Where America stands: Education in America. (2010): Retrieved from http://www.cbsnews.com/htdocs/pdf/poll_whereamericastands_education_012110.pdf
- Chukwubikem, E.I. (2012). Developing better teacher evaluation. *International Journal of Social Sciences & Education*, 2(4), 554-568.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity: Summary report* (Vol. 2). US Department of Health, Education, and Welfare, Office of Education.
- Conley, D.T. (1987). Critical Attributes of Effective Evaluation Systems. *Educational Leadership*, 44(7), 60-64.
- Corcoran, S. P. (2010). Can Teachers Be Evaluated by Their Students' Test Scores? Should They Be? The Use of Value-Added Measures of Teacher Effectiveness in Policy and Practice. Education Policy for Action Series. *Annenberg Institute for School Reform at Brown University* (NJ1).
- Creswell, J. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd ed.). Thousand Oaks, CA: SAGE Publications.
- Delvaux, E., Vanhoof, J., Tuytens, M., Vekeman, E., Devos, G., Petegem, P.V. (2013). How may teacher evaluation have an impact on professional development? A multilevel analysis. *Teaching and Teacher Evaluation*(36), 1-11.
- Hiller, J. (1986). Teacher evaluation: Issues, practices, and a semiotic model. *Education*, 107(2), 144-148.
- Jackson, C.K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, 1(4), 85-108.

- Kane, T.J., Taylor, E.S., Tyler, J.H., & Wooten, A.L. (2011). Identifying effective classroom practices using student achievement data. *Journal of Human Resources*, 46(3), 587-613.
- Kersting, N.B., Chen, M.K., Stigler, J.W. (2013). Value-added teacher estimates as part of teacher evaluations: exploring the effects of data and model specifications on the stability of teacher value-added scores. *Education Policy Analysis Archives*, 21(7), 1-40.
- Koedel, C. (2009). An empirical analysis of teacher spillover effects in secondary school. *Economics of Education Review*, 28(6), 682-692.
- Looney, J. (2011). Developing High-Quality Teachers: Teacher evaluation for improvement. *European Journal of Education*, 46(4), 440-455.
- Maslow, V.J., & Kelley, C.J. (2012). Does Evaluation Advance Teaching Practice? The Effects of Performance Evaluation on Teaching Quality and System Change in Large Diverse High Schools. *Journal of School Leadership*, 22(3), 600-632.
- Mathers, C., & Olivia, M. (2008). Improving Instruction through Effective Teacher Evaluation: Options for States and Districts. Retrieved from <http://files.eric.ed.gov/fulltext/ED520778.pdf>
- Minnesota Department of Education (2014). Data Center. *Minnesota Department of Education*. Retrieved April 29, 2014, from <http://www.education.state.mn.us>
- Olivia, M., Mathers, C., & Laine, S. (2009). Effective evaluation. *Principal Leadership*, 9(7), 16-21.
- Paige, M. (2013). Applying the Paradox Theory: A Law and Policy Analysis of Collective Bargaining Rights and Teacher Evaluation Reform from Selected States. *BYU Education and Law Journal*, 21.
- Stark, J.S. & Lowther, M.A. (1984). Predictors of teachers' preferences concerning their evaluation. *Educational Administration Quarterly*, 20, 76-106.
- Taylor, E.S., and Tyler, J.H. (2012). Can teacher evaluation improve teaching? *Education Next*, 12(4), 79-84.
- Teachers and Other Educators, Minnesota Statute § 122A.40 (2014).
- Tompson, T., Benz, J., & Agiesta, J. (2013). Parents' attitudes on the quality of education in the United States. *Associated Press – NORC Center for Public Affairs Research*.
- US Department of Education (2014, August 20). National Assessment of Educational Progress. National Center for Education Statistics. Retrieved September 6, 2014, from <http://www.nces.ed.gov/nationsreportcard>
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. Retrieved from http://tntp.org/assets/documents/TheWidgetEffect_2nd_ed.pdf
- Zakariya, S.B. (1985). Rand Corp.: These Are the Elements of a Sound Teacher Evaluation System. *American School Board Journal*, 172(7), 25.

